DOCUMENT RESUME

ED 151 682 CG 012 283

AUTHOR MacKinnon, Donald W.
TITLE An Overview of Assessment Centers. Technical Report 1.
INSTITUTION Center for Creative Leadership, Greensboro, N.C.
PUB DATE May 75
NOTE 46p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Ethics; *Evaluation; Growth Patterns; Historical Reviews; *Industry; *Management Development; *Psychoeducational Processes; *Staff Improvement; State of the Art Reviews; *Success Factors
IDENTIFIERS *Assessment Centers

ABSTRACT

This overview examines the historical growth of industrial assessment centers; describes the techniques involved in assessment; discusses the role of management in staff development; and suggests new uses for assessment centers. (CKJ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151682

An Overview Of Assessment Centers

Donald W. MacKinnon

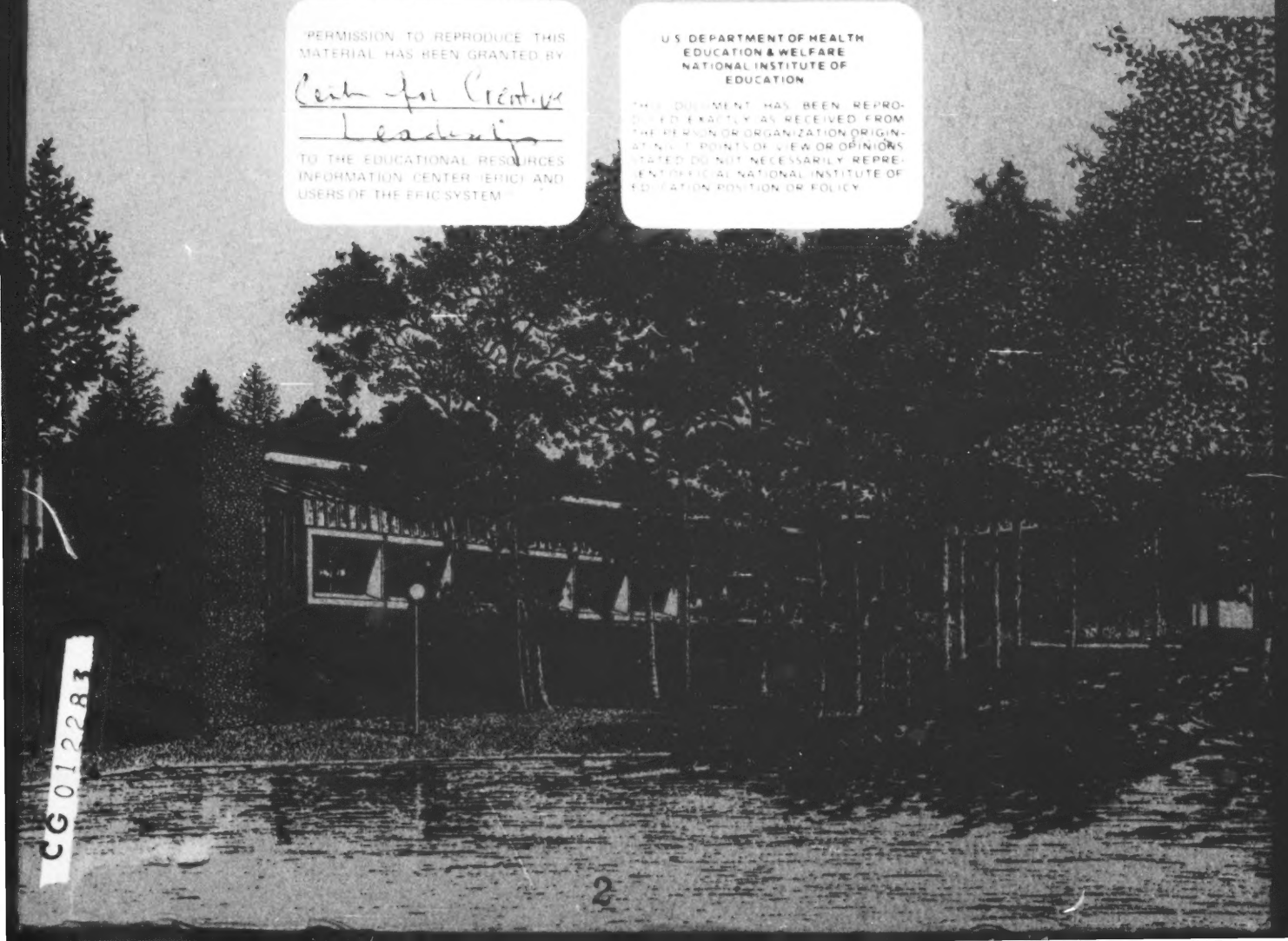
Technical Report No. 1

CENTER FOR CREATIVE LEADERSHIP

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Center for Creative Leadership
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.



CG 012281

AN OVERVIEW OF ASSESSMENT CENTERS

Donald W. MacKinnon, Ph.D.
University of California, Berkeley

CCL Technical Report
No. 1
May, 1975

An Overview of Assessment Centers[1]

Donald W. MacKinnon

University of California, Berkeley

I hardly need tell industrial psychologists that assessment is a method for the psychological evaluation of individuals that involves testing and observing of individuals in a group setting, with a multiplicity of tests and procedures, by a number of staff members. Through a pooling of test scores and subjective impressions, the assessors formulate psychodynamic descriptions of the assessed subjects which, hopefully, will permit prediction of the assessees' behavior in certain kinds of roles and situations.

The first assessment center in the United States, Station S in the Office of Strategic Services, was set up to evaluate the qualification of candidates to serve overseas in irregular warfare during World War II. Its methods and procedures owed much to the pioneering efforts of German military psychologists during the 1930s, most notably the work of Simoneit (1940, 1944), and to the programs developed by the War Office Selection Boards (WOSBs) for the testing of officer candidates for the British Army (Garforth, 1945).

The philosophy and theory of assessment as it was developed in the OSS derived most importantly from the work of Henry A. Murray (1938) and his associates in the Harvard

---

[1]Adapted from a paper presented at the Industrial Psychologists Meeting, Center for Creative Leadership, Greensboro, North Carolina, January 29-30, 1975. A briefer version entitled, "The Measure of Man," is being published as an occasional colloquium paper by the Department of Psychology, Lakehead University, Thunder Bay, Ontario, Canada.

I want to thank the Center for Creative Leadership for making it possible for me to survey assessment centers during my year as Visiting Fellow at the Center in 1973-74. Thanks are also due to those directors of assessment programs who permitted me to observe their centers and who generously shared their knowledge and experience with me. Without both of these sources of help, this report could not have been written.

Psychological Clinic during the '30s. Emphases in the OSS assessment program, which had been largely lacking in earlier programs, were 1) an attempt to delineate the person as a whole, 2) a concern with the more favorable aspects of personality and its potentialities for effective functioning in demanding roles and situations, and 3) a serious attempt to test the validity of the predictions concerning the future performance of assessed subjects.

The early reports on the OSS assessment program (Murray & MacKinnon, 1946; MacKinnon, 1946, 1947) stirred considerable interest in the use of the assessment method, not only for predicting future performance but also as a research method for the intensive study of highly effective individuals. In 1946 an assessment study of clinical psychologists was undertaken at the University of Michigan (Kelly & Fiske, 1950, 1951; Kelly & Goldberg, 1959), and in 1947 a similar study of psychiatrists was launched at the Menninger Foundation (Holt & Luborsky, 1958), and in 1949 the Institute of Personality Assessment and Research was established at the University of California, Berkeley, as a continuing assessment center (MacKinnon, 1960).

When the final report on the OSS program, Assessment of Men, appeared in 1948, some of the earlier enthusiasm for the assessment method was dampened. The OSS staff had taken seriously its obligation to check the validity of their operations but for a number of reasons, all of which were spelled out in their book, they were unable to do so very effectively. Nor did the other immediately post-war assessment programs fare very much better, and--prematurely, as it turned out--Cronbach (1955,1956) declared that the OSS style of assessment had failed. He was, like other critics of assessment programs, preoccupied with the relatively low validities of several early assessment studies and overlooked other equally important goals of these programs; namely, the objective of discovering what goes with what in the personality structure of individuals who have the capacity to realize their potential for growth and development or who have already become extraordinary human beings either in their personal lives or professional careers.

Had Cronbach been more prescient, he would have foreseen that just one year later, 1956, Douglas W. Bray would revitalize the assessment method in the business world by studying a group of beginning managers in AT&T and following their development in a major longitudinal research, the now famous Management Progress Study (Bray, Campbell, & Grant, 1974).

If Bray was to discover the changes that occur as the result of following a career in business management, he had first to determine the baseline personalities of his subjects

at the initiation of his research. Of the techniques available for the study of these persons in all their complexity, it seemed to him that the assessment method was the most promising. Thus, the first assessment center in American industry was established in 1956, not for operational goals of selection and placement, but for research purposes.

Growth of Assessment Centers

This first assessment center was set up in the Michigan Bell Telephone Company where it generated a good deal of interest, and two years later, 1958, Michigan Bell became the first of the Bell System companies to establish an assessment center program for operational use, specifically to assess the qualifications of relatively long-term non-management employees to move into managerial positions. Soon thereafter, centers for the assessment of managerial potential were set up in Standard Oil of Ohio, IBM, Sears Roebuck, General Electric, Wickes, and J. C. Penney. Although some of these are no longer in existence, others in increasing numbers have come to take their place. In reading the literature and in conversation with those who ought to know, one could estimate that the number of assessment centers today ranges from 300 to more than 1000.

Why such a range? It is partially due to the rapidity with which new assessment centers are now being established, and partly due to some confusion as to what an "assessment center" is. The term sometimes refers to a physical facility, but there are few assessment centers in this sense, one being the Institute of Personality Assessment and Research at Berkeley, another being the Center for Creative Leadership where we meet today. Sometimes an "assessment center" is taken to be a program of assessment such as FACT (the Formal Assessment of Corporate Talent) in Standard Oil of Ohio, or AT&T's MAP (Management Assessment Program). Thusly defined, the number of assessment centers is the number of companies that run assessment programs, perhaps 300 or more. There is a third sense in which the term "assessment center" is used; namely, to designate the coming together of a group of candidates to be assessed. In this sense there were, for example, more than 50 assessment centers conducted in the Office Products Division of IBM in 1974. This appears to be the most common meaning of "assessment center" and leads to the higher estimate of their number.

Why has there been such a burgeoning of interest in assessment centers in the '70s? There are several factors contributing to what can only be called the assessment center movement. One has been the increasing need for

effective managers and the realization of how costly it is to hire managers who fail. Firm data are hard to come by, but a study by Syntex (Miller, 1974) has come up with an estimate that hiring a poor sales manager rather than an average one, to say nothing of a superior one, can cost as much as $75,000 a year. Whether this figure is realistic or not, it is widely believed that hiring ineffective managers can be not only costly for a company but, in certain key positions, disastrous. This increases the interest in utilizing any method of selecting managers that promises to be more efficient than traditional ones. Such a method is the assessment center.

Earlier programs of psychological testing and interviewing by a psychologist or psychiatrist have not demonstrated their effectiveness on any large scale. The individual psychologist offering his services in finding potential managers is often perceived (whether correctly or not is beside the point) as claiming certain esoteric skills and conveying an aura of mystique as he goes about his business. Thus he draws a certain amount of resistance and rejection.

Assessment centers, on the other hand, are less mysterious, more open, and--above all--have high face validity. In situations which simulate those that industrial managers face every day, candidates are confronted with demands, pressures, and problems, and are asked to deal with them. Their coping styles are observed, not by a single evaluator but by a staff of observers, often including managers themselves, and the relative effectiveness of the candidates' performances are noted. If participants react effectively to those simulated tasks, the philosophy of assessment assumes that they will perform with similar effectiveness if promoted to managerial positions. The basic assumption is that the best predictor of future behavior is past behavior. To many, this seems both plausible and fair, and is one of the reasons why assessment centers have won such widespread acceptance.

While the face validity of assessment centers may lead to a favorable first impression, that is not sufficient. Usually, more solid evidence of their validity is required before companies are ready to make the heavy investment that the setting up of an assessment center entails. For this reason, a key factor in the increasing popularity of assessment centers has been the Management Progress Study in AT&T (Bray et al., 1974).

The history of assessment centers in industry would, I believe, have been quite different had the first center been set up for operational use rather than as a research effort not only to discover what happens in the lives of managers but even more importantly to determine the validity of

assessment centers. The positive and impressive findings of this study have been crucial. Indeed, it has been asserted that on the basis of this single study, "the assessment center has made its reputation in American industry" (Cohen, 1974a, p. 19). One should not, however, overlook the influence that other research studies of assessment centers have had in creating a climate of acceptance of assessment programs in industry.

The Civil Rights Act of 1964, leading to the Equal Employment Opportunity Commission Guidelines published in 1966 and revised in 1970, has played a significant role in the increasing use of assessment centers. The EEOC Guidelines have been widely interpreted as banning, or at least rejecting, the use of psychometric tests in employee selection. Tests have not been prohibited. Rather, those tests used in making employment decisions must be relevant to success on the job, and they must not discriminate against members of minority groups. Since unequivocal evidence of this kind exists for few tests, there has been tremendous pressure to substitute safer and fairer methods of personnel selection. Among these methods, assessment centers have acquired a special attractiveness, especially since Huck (1973), in a Michigan Bell study, showed that women and blacks performed as well as men and whites in assessment center exercises.

But possibly the factor that has made for the most rapid multiplication of assessment centers has been the establishment of managerial consulting companies that are prepared to establish assessment centers in any industrial firm which might desire them, both here and abroad. The two major suppliers of such complete assessment centers set up within individual companies are the American Management Association (McConnell & Parker, 1972), which offers such services to its client members, and Development Dimensions, Inc., which offers its services to any interested, qualified company. The former was set up by McConnell (Cutterbuck, 1974), the latter by Bray and Byham. Among the services offered by Development Dimensions (1975) are aid in starting an assessment center, custom-made as well as standard assessment exercises, evaluation of on-going assessment centers, and both administrators and professional assessors for an organization's assessment center. In addition, it schedules each year a number of orientation conferences and workshops on the use of assessment centers and assessment center methods.

There are doubtless other factors that are responsible for the phenomenal growth of assessment centers, such as, for example, Byham's (1970) seminal paper in The Harvard Business Review. But for whatever reasons, the assessment

center concept is well established and widely accepted in industry today.

Variations on a Theme

To describe each of the major assessment centers would make for a colorful and fascinating report, but there is not time for that here. Instead, to convey some idea of the differences among them, I propose to sketch the major dimensions on which assessment centers vary. To be sure, all assessment centers have much in common. If this were not so they would not be assessment centers. But no two of them are exactly alike. They are, as Erwin Taylor (Finkle, Note 1) has suggested, "variations on a theme." Among the major dimensions on which assessment centers vary are these:

Methods and Techniques of Assessment

Few assessment programs in industry use exactly the same exercises, but all of them place heavy reliance upon the observation of behavior in contrived situations, designed to simulate the situations that business managers frequently experience. Typical procedures included in an assessment battery are:

an In-Basket which must be quickly completed;

a Leaderless Group Discussion of a Cooperative Nature in which, for example, a candidate assumes the role of a business consultant who--with five other consultants--receives data on a particular company and has the task of helping the group to decide whether to recommend that the company continue to operate or to sell out;

a Leaderless Group Discussion of a Competitive Nature in which, for example, each candidate is assigned the role of vice-president (Public Relations, Personnel, Manufacturing, and the like) of a company which, in a period of expansion, is considering the purchase of one of three new sites. The six "vice-presidents" are meeting to decide which of the sites should be selected. The conflict for each assessee is obvious: to win selection of the site he favors vs. selection of the best possible site for the company.

Dyadic situations are also employed in which, for example, the assessee is interviewed by his supervisor (played by a member of the assessment staff) concerning the actions he took in clearing his In-Basket; or the assessee, (playing the role of a manager), interviews a capable but troublesome employee, a standardized role played by a staff member.

Individual exercises are frequently also included, such as a Writing Exercise as an example of written communication skills or a Speaking Exercise as an example of oral communication skills.

Other procedures sometimes used are a Life-History Interview with the assessee, which frequently builds upon a Biographical Inventory Blank which the candidate has filled out. The interview typically covers an assessee's personal history, work history, and the history of his goals and values. Peer Ratings and Sociometric Ratings as well as Self-Ratings are often collected. Projective Tests, most often the Thematic Apperception Test and a Sentence Completion Test, may be used to reveal the quality and strengths of the assessee's motivations, especially for a career in management.

Finally, Psychometric Tests of mental ability, interests, values, and personality are administered in some centers, but typically test scores are not used in making decisions or recommendations concerning candidates. They are used as a check upon decisions or recommendations already agreed upon, and are sometimes retained for later research. The time may come when psychometric tests may indeed be demonstrated to be efficient predictors of managerial behavior.

Assessment centers bring assessees and assessors together for varying periods of time. During that period, at meals, at coffee breaks, and at cocktail hours (if they are included in the program), a large variety of behaviors may be observed. Some assessment centers believe all these casual observations are grist for the assessment mill and they are so used; others believe that assessments and recommendations should be based solely upon behaviors and responses revealed in the formal exercises.

Number of Tests and Procedures Used in Assessment Centers

The number of evaluative procedures used in assessment centers has been reported as ranging from 4 to 40, the modal number being 7 (Bender, 1973).

Length of Assessment Centers

There is, of course, a relationship between the length of an assessment center and the number of assessment procedures employed. Assessment centers vary in length from one day to one week.

Dimensions of Behavior Rated by Assessors

The number of variables or factors thought to be related to managerial performance and rated by assessment staffs has been reported as ranging from 10 to 52 (Howard, 1974). In the assessment centers which I have personally observed, the range was from 10 to 38.

Administrative Arrangements

Assessment centers are either set up within the companies they serve or by management consulting firms as a service to any companies that may wish to use them.

Staffing

Historically, assessment centers have been developed by psychologists. They are sometimes staffed by psychologists, sometimes by members of management--typically two or three levels above the assessees--and sometimes both psychologists and managers serve as assessors. The person who directs the assessment center and serves as chairman of the evaluation session in which assessment findings are integrated may be either a member of management or a psychologist, and may play an active role in the discussion or serve only as chairman of the meeting.

Length of Service as Assessors

Length of service as assessors may extend over a period of years, as in the case of professional psychologists or personnel officers; or last six or seven months, in the case of second- or third-level managers who are assigned to an assessment center; or in the case of higher level managers (such as Vice-Presidents) be brief, usually a one-time event.

Training of Assessors

Training of assessors may be limited, involving only a

general orientation with brief instructions before each exercise. In some cases, the variables to be rated are not discussed until the integration session. At the other extreme, training may extend over years of study in clinical and industrial psychology leading to the Ph.D., followed by detailed training as an assessor. When managers serve as assessors over a period of months, training usually continues over one to three weeks. During this time they may experience assessment as an assessee, watch videotapes of assessment exercises, study assessment manuals and forms, serve as an observer of an assessment center making ratings as non-voting members of the staff, and have their assessment judgments checked by an experienced assessor.

Number of Assessors on a Staff

The number of assessors serving on a staff varies from 3 to 24, with 6 being the most typical staff size (Bender, 1973).

Number of Assessees Evaluated in an Assessment Center

The number of assessees evaluated is typically six or a multiple thereof, since six has been found to be an optimal number of assessees to interact and to be observed in simulation exercises. The numbers actually assessed in assessment centers has been reported as 6, 8, 10, 12, 18, 24, and 84, 6 being the mode (Bender, 1973).

Ratio of Assessors to Assessees

The ratio of assessors to assessees is usually about one to two or three.

Selection of Candidates for Assessment

Those selected to be assessed have typically demonstrated superior performance in non-management roles or in managerial positions and are judged to have potential to perform effectively at some higher level of responsibility. Usually they are nominated by a supervisor, sometimes with and sometimes without approval by higher levels of management. In some companies, self-nominations are permitted--a practice initiated in Michigan Bell as early as 1964. Participation in an assessment center is voluntary and can be refused without prejudice to the nominee. This, at least, is the

usual policy statement, but pressure to attend, if nominated, is frequently felt, and the consequences of refusing to attend are not infrequently feared.

Purpose of Assessment Centers in Industry and Business

Broadly speaking, the purpose of assessment centers in business and industry is to identify managerial potential for purposes of selection, placement, promotion, or development (or some combination of these), but which of these is most stressed varies from program to program.

Levels of Assessment

The levels at which candidates are assessed vary, as do the outcomes for which assessments are made: for example, college graduates for hiring, the newly-hired for accelerated training, the experienced worker for first-line management, first-line managers for middle management, middle-level managers for upper-level management, and upper-level managers for top-level executive.

Integration of Assessment Findings in Staff Meeting

The products of assessment centers--feedback to the assessees and reports to the company--are prepared in a staff meeting in which assessment findings are evaluated and integrated, with reports being written after the session. The length of these meetings, may be as brief as 1½ hours or extend over a period of a day or two.

Feedback to Assessees

With the exception of assessment centers which are conducted for pure research, such as in the AT&T Management Progress Study, feedback from the assessment center is usually given to those who have been assessed. The form may be an oral report or an oral report combined with a written report, which may or may not be left with the candidate. In some cases, a tape of the feedback interview is provided to the assessee so that he may review it later.

The content of the feedback usually includes a description of the candidate's performance in the assessment center, an indication as to how he was viewed by his peers as well as by the assessment staff, a review of his judged strengths and weaknesses, and in some instances, an outlined

plan for his future development as a manager or even as a person.

The giver of the feedback is usually the staff member who wrote the report on the candidate. This is sometimes the director of the center or chairman of the integration session; sometimes it is the assessor who conducts the life history interview; and sometimes it is the psychologist, whose task it is to render as accurate a report as possible of the consensus of the assessors as expressed in the final integration session.

The place where the feedback is given may be the assessment center itself, but more often it is some place mutually convenient for the assessee and the feedback-giver and is frequently away from the candidate's place of work.

The time of giving the feedback may be at the end of the assessment period or weeks later.

The person or persons to whom the feedback is given also varies. The feedback is prepared for the individual assessee and is intended to be helpful to him. Usually it is for the candidate alone, but sometimes it is shared with his supervisor.

Report to Management

In addition to the feedback to assessees, written reports are prepared for company records. (An exception would be centers conducted for pure research.) These are usually written shortly after the integration sessions and are supposed to reflect the consensus of the assessment staff. In other assessment programs, the report is not written until after the feedback has been given to the candidate. Where this is the practice, the feedback interview provides further information about the candidate which is incorporated into the report.

Just as the feedback is designed primarily for the candidate, although sometimes shared with his supervisor and/or manager, assessment reports are prepared for management--usually for managers at least two or three levels above the candidate--and for the personnel division.

In general, reports are of one or two kinds, depending upon the purpose of the assessment program. Most typically they describe the performance of the candidate in the assessment center, and provide rating of the candidate on the rated variables, with an overall rating of the candidate's managerial potential if selected or promoted. In companies

that use the assessment findings in creating tailor-made programs for the development of their personnel--both as persons and as managers--a second report, or a second part of the single report, is concerned with developmental recommendations.

Practice varies in the disposition of assessment center reports. Generally, they are not put in the individual's personal folder but filed with the assessment center staff and with the personnel division headquarters. How long the reports should be retained is an important question answered differently by different companies.

Why are there so many variations in practically all aspects of assessment centers? In large measure, the variations are the result of ignorance. There are both theoretical reasons and practical considerations underlying each of the variations, but in the absence of research designed to determine which variations lead to more effective--which is to say, more valid--assessments, the selection of one rather than another variation by any assessment center remains largely a matter of personal preference. What is needed is research, and research on a grand scale, to determine which of the practices are more effective and thus to be preferred and under what circumstances.

To mention only a few possibilities, the relative merits of a one-day assessment versus an assessment of two or three days could be investigated. So also could the advantages and disadvantages of an assessment program composed of exercises unrelated to each other in terms of their content be compared to those of a program in which the exercises are so integrated that they simulate an industrial or business unit.

Another research task might be to determine the optimal timing of feedback, whether it should be given during the assessment center or at some later date, and if later, at what time, and in what place.

I do not wish to imply, by these remarks, that no methodological research has been conducted. There has been some, but not nearly enough, and by only a small number of assessment centers. What has been demonstrated for them does not necessarily hold for the hundreds of assessment centers whose operations have not been critically examined.

In setting up the assessment center for the Management Progress Study, Bray (Bray & Grant, 1966; Bray et al., 1974) compiled a list of 25 qualities or attributes thought to be

important for managerial success in AT&T. They were eventually the variables rated in the assessment center:

1. Scholastic Aptitude
2. Oral Communication Skill
   Written Communication Skill
3. Human Relations Skills
4. Personal Impact
5. Perception of Threshold Social Cues
6. Creativity
7. Self-Objectivity
8. Social Objectivity
9. Behavior Flexibility
10. Need Approval of Superiors
11. Need Approval of Peers
12. Inner Work Standards
13. Need Advancement
14. Need Security
15. Goal Flexibility
16. Primacy of Work
17. Bell System Value Orientation
18. Realism of Expectations
19. Tolerance of Uncertainty
20. Ability to Delay Gratification
21. Resistance to Stress
22. Range of Interests
23. Energy
24. Organization and Planning
25. Decision Making

To these 25 variables, a 26th was added, "Overall Staff Prediction." This is a simple staff rating of managerial potential; that is, the level that an assessee is predicted to reach during his career in the company.

Reliability

The pattern set by Bray of conceptualizing a fairly large number of variables has been followed by most assessment centers. The question arises as to how reliably participants can be rated or ranked on these dimensions.

Several studies support quite clearly the conclusion that inter-rater reliabilities in assessment evaluations are sufficiently high to justify their further use. For example, in the Management Progress Study (Bray & Grant, 1966), the inter-rater reliabilities (for two psychologists) were, for ratings and rankings based on observation of a leaderless group discussion, both .75; based on observation of a manufacturing problem, .60 and .69; while ratings based only on an In-Basket had a reliability of .92.

In the IBM program (Greenwood & McNamara, 1967), the inter-rater reliabilities for ratings and for rankings for all three pairs of alternating observers were, based on a task force game, .70 and .71; based on a leaderless group discussion, .66 and .64; and based on a manufacturing problem, .74 and .75. The median of the 432 reliability coefficients computed for ratings was .74, and for rankings .76. When the inter-rater reliabilities were computed separately for various pair-combinations of psychologists and line managers, no consistent differences in degree of agreement between particular types of pairs of assessors were found. That is to say, psychologist agreed with psychologist to about the same degree that psychologist agreed with manager, and manager with manager, in rating the assessment variables. Further, means and standard deviations between the ratings by psychologists and managers were not significantly different.

Similar findings were made in the SOHIO Program (Thomson, 1969;1970). Inter-rater reliabilities of ratings by two psychologists on 13 dimensions ranged from .73 to .93, $\bar{r}$ =.85. When the ratings were made by three managers, the inter-reliabilities ranged from .78 to .95, $\bar{r}$=.89. Remarkable agreement between psychologists' and managers' ratings was also found. Inter-correlations of their ratings on the 13

SOHIO rating dimensions ranged from .74 to .93, as follows:

Rating Scale	Correlation
Amount of Participation	.93
Oral Communication	.85
Personal Acceptability	.80
Impact	.89
Quality of Participation	.81
Personal Breadth	.82
Orientation to Detail	.76
Self-Direction	.86
Relationship with Authority	.84
Originality	.85
Understanding People	.80
Drive	.74
Potential	.91

At least in those centers that have researched their operations, behavior observed during an assessment center can be recorded and rated with high reliability by staff observers (whether psychologists or line managers) and ratings and rankings can be made with about the same reliability.

Factor Analytic Studies

With so many variables being rated, the redundancy among them must be considerable. The question arises, then, as to just what dimensions of behavior are indeed discriminated by staff observers in their ratings.

To answer this question, factor analytic studies of assessment center ratings have been conducted in four major assessment programs: The AT&T Management Progress Study (Bray, 1964a, 1964b; Bray & Grant, 1966; Bray et al., 1974); SOHIO's Formal Analysis of Corporate Talent (FACT) program (Thomson, 1969; Donaldson, 1969; Finkle & Jones, 1970), the IBM program (Hinrichs, 1969; Hinrichs & Haanperä, Note 2; Kraut, 1969, 1972a, 1972b; Kraut & Scott, 1972), and the Sears program (Bentz, 1967, chap. 7; Bentz, Note 3). Beyond yielding certain factors dependent upon the particular variables rated in each program, these four studies are summarized by Dunnette (1971) as showing the following broad

underlying dimensions:

Overall Activity and General Effectiveness

Organizing and Planning (Administering)

Interpersonal Competence (Human Relations Ability and Understanding)

Cognitive Competence (Intellectual and Quality of Thinking)

Work Orientation (Work-Oriented Motivation)

Personal Control (Control of Feelings and Resistance to stress)

Despite the redundancy among rated variables, many assessment centers continue to rate large numbers of them rather than the smaller number of rating factors. This seems to me not a bad strategy, for it forces staff members to strive for differentiated perceptions of their assessees' behavior. If, as is the case, the conceptualized dimensions can be observed and rated with high inter-rater agreement--a spread of ratings greater than one step is infrequent--meaningful distinctions are being made which may be clinically if not statistically justified.

The question of which assessment procedures contribute most importantly to behavior rating factors, as well as to overall staff predictions, is one which must be answered separately for each assessment center. A good illustration of one answer can be found in the Management Progress Study college sample (Bray et al., 1974). Although the analysis is not yet completed, it is pretty clear that the procedures which contributed most importantly to the rating of each of the seven rating factors were these:

In-Basket	to Administrative Skills
Group Exercises	to Interpersonal Skills
Ability Tests	to Intellectual Ability
Simulations	to Stability of Performance
Projective Tests & Interview	to Work Motivation
Projective Tests & Interview	to Career Orientation
Projective Tests	to Dependency on Others

And the procedures which contributed most significantly to the overall staff prediction were Group Exercises and the In-Basket.

Validity

An assessment method designed to reveal and measure human effectiveness, both real and potential, may well be asked to demonstrate its own effectiveness. The validity of the assessment method has been extensively researched, but by only a small number of assessment centers if the count is based upon published researches. A variety of criteria of success has been employed against which the predictive validities of component elements of the assessment method, as well as overall ratings and predictions, have been measured; namely (1) job performance (ratings of performance on the job for which the candidate was assessed, (2) job progress (promotions, demotions, increases in salary, etc.); and (3) job potential (ratings of the likelihood of future progress). Time permits only the merest sketching of some typical findings.

The study involving the largest number of subjects included 5943 men assessed in AT&T (Moses, Note 4) and placed in one of four rating categories ranging from "more than acceptable" to "not acceptable." The question was: what percentage of assessees in each of the categories met the stringent criterion of having received two or more promotions since assessment? This criterion was met by 41% of those rated "more than acceptable," by 22% of those rated "acceptable," by only 12% of those rated "questionable," and by 4% of the "not acceptable." The correlation between overall assessment rating and progress was .44($p<.001$).

In an earlier IBM study (Wollowick & McNamara, 1969), the correlation of the overall assessment rating with increase in managerial responsibility (taking account of such factors as numbers of persons supervised, job complexity, skill and knowledge required, fiscal responsibility, etc.) over a three year period was .37 ($p<.001$). In another IBM study, the predictive validity of ratings for 11 assessment groups over periods of one to four years was examined and significant correlations were found for 8 of the 11 groups with r's ranging from .29 to .63 (Dodd, Note 5).

In the SOHIO FACT program, significant validities have also been found of .64 ($p<.001$) between overall ratings and ratings based on supervisors' interviews 6 - 27 months after assessment (Thomson, 1970); when the criterion was supervisors'

rating of managerial potential made 9 - 29 months after assessment, the predictive validity of the overall rating was .63 ($p<.001$); when made 30 - 62 months after assessment, it was .65 ($p<.001$) (Finley, 1970); and in another study with ratings made 2½ - 5 years after assessment, the correlation was again .65 (<p.001) (Carleton, 1970).

Several studies of the validity of the component parts of the assessment armamentarium have also been made. Especially comprehensive is the one reported for the Sears program (Bentz, Notes 6, 7, 8, 9). Assessment ratings based on tests, on an In-Basket, and on group exercises have been correlated with on-the-job performance ratings made by personnel directors, by the candidate's supervisors, and by the candidate himself, as well as with such progress criteria as indices of job mobility and of salary progress. Pages of significant validity coefficients have been presented in one report alone, some type of validity having been shown for all components of the Sears program (Bentz, Note 10).

Only one study in the U.S. has reached a negative judgment on the assessment method. At the Caterpillar Tractor Company, the performance of 37 men appointed as first-line shop supervisors following assessment was compared with that of 27 men whose appointment was based on the traditional method of selection. The finding: supervisors selected by the traditional method were higher in job performance than those picked by the assessment centers (Bullard, Note 11).

The failure of the assessment method here may have been related to the selection of the samples: only "obviously" superior candidates were promoted without assessment, while those assessed were marginal performers. Under these circumstances, the representativeness of the two samples may well be questioned. Or the failure may have resulted from the technical demands of the shop supervisor's position mounting larger than managerial demands in determining successful performance, with technical skills not being evaluated in the assessment program (Cohen, Moses & Byham, 1974).

There is, however, one shortcoming in these validity studies, just as there is in all the validity studies found in the literature, save two; namely, that one cannot know to what extent the criteria have been contaminated by the predictors, for these researches have been made in operational programs where line management has had the benefit of the assessment center reports.

It has generally been assumed that whatever error has resulted from this has spuriously raised the validities, but

how much, if indeed at all, one cannot say with certainty. For this reason, the two pure validity studies--in which assessment data have in no way influenced personnel decisions--have been crucial for the expanding use of assessment centers, and both have been made in AT&T where it all started.

The first of these, indeed "what must be regarded as 'The Study' in assessment center validity" (Howard, 1974, p. 122) is the Management Progress Study (Bray & Grant, 1966). Two samples (one of college men and one of non-college men embarking upon careers in management) were intensively assessed, with no assessment data ever being made available to management or to the participants. The major staff prediction for each man was that he would or would not make middle management within 10 years of the time of assessment.

These predictions were checked 5 to 8 years later against the level of management actually achieved. For both samples combined, of those predicted to make middle management 42% did so, while only 7% of those predicted to remain at the entry level were promoted to middle management. Conversely, of those who were predicted to reach middle management only 4% remained at the first level, while 42% of those predicted not to move above the entry level of management remained there.

Expressing the results somewhat differently and separately for the two samples: in the college sample, 31 (82%) of the 38 men who made middle management had been correctly identified. Again combining the two samples, 68 (94%) of the 72 men who did not advance beyond the first level were correctly identified by the assessment staffs (Dunnette, 1971).

The other pure validity study (Bray & Campbell, 1968), also conducted in AT&T, involved the assessment of candidates for sales rather than managerial skills. Without any report back to the participants or to the company of assessment findings, all who had been assessed were hired as salesmen. Six months later their performances were evaluated by a team of experienced sales reviewers. Four categories of assessment were employed, ranging from "More than Acceptable" to "Unacceptable," and the percentage in each category who met the review standards were 100% of the "More than Acceptable," 60% of the "Acceptable," 44% of the "Less than Acceptable," and 10% of the "Unacceptable." The chi-square value was 24.19 ($p<.001$).

The findings of both the pure and the operational validity research studies lead to the conclusion that assessment centers can be remarkably effective in identifying managerial potential--but precisely how effective, one

cannot say. Generally, it has been assumed, because of the possibility of contamination of the criteria by the predictors, that validities have been spuriously high. There are, however, also reasons to think that in some instances they may have been spuriously low due to invalidity of the criteria. The high percentage of unqualified among candidates nominated by supervisors for assessment suggests that supervisors' judgments are far from infallible and their ratings of job performance questionable.

Also, as already noted, managers serving as assessors in the FACT program (Thomson, 1970) showed high agreement with psychologists in rating candidates on 13 dimensions of managerial behavior (median r =.85 between raters). But when supervisors were asked to rate these same persons on these same dimensions some 6 to 24 months later, their ratings were less reliable, more restricted in range, and more skewed on the side of leniency than the ratings made by the manager-assessors. Surely, such findings must call into question the appropriateness of using supervisors' ratings as criteria of job performance.

This is not to suggest that managers cannot be effective evaluators or judges but rather that they are unlikely to make finely differentiated evaluations unless they receive the same kind of experience or training that managers who serve as assessors in assessment centers receive.

As Cohen, et al.,(1974) have observed, "Since the same supervisors who nominated candidates are often called on to evaluate subsequent performance, the uncritical use of supervisory judgment as the sole basis of validating assessment performance is a questionable procedure. As noted by Bray and Moses (1972), it is somewhat absurd to validate an expensive, high-powered (in the sense of information generated) process by a cheap, often unrealistic but easily obtainable 'criterion' measure" (p. 21).

Given all the uncertainties in the research on the validity of assessment centers, I still think it prudent to conclude that the best of them are remarkably effective in identifying managerial potential. But even if high validity could be proved beyond peradventure, that would not suffice. Assessment centers are a costly way of finding managerial talent. Accordingly, they must be demonstrably more valid than the traditional, less expensive methods of selecting managers if they are to be fully justified. Such demonstrations are not lacking.

An early study in Michigan Bell (Note 12) compared the on-the-job performance of the last 40 men promoted to management before the assessment program was initiated with the

first 40 men who were promoted on the basis of assessment center evaluations. While only 35% of those not assessed were rated "better than satisfactory" in job performance, 63% of those assessed and then promoted received a "better than satisfactory" rating. Further, while only 35% of the non-assessed group were rated as having the abilities required for the next level of management, 67% of the assessed sample were so rated. These findings stand not alone but have been supported by similar studies in other companies (Campbell & Bray, 1967; Jaffee, Bender & Calvert, 1970).

A study in one General Electric plant (Meyer, Note 13) compared the demotion rate of 36 foremen promoted without benefit of assessment with a roughly-matched group of foremen whose promotions had been based partly on assessment data. Six of the non-assessed foremen were subsequently removed for poor performance, but none who had been assessed were demoted.

With the exception of the study in the Caterpillar Tractor Company (Bullard, Note 14), when the performance of men promoted into management positions without assessment has been compared with that of men promoted after having been assessed, the latter have been shown to be more successful. Cohen, et al.,(1974) have tabulated the median percentages of success for both assessed and non-assessed groups and have found the assessed group higher in job performance, .68 vs. .59; higher in job potential, .59 vs. .28; and higher in progress, .38 vs. .08. Taken at face value, these figures would appear to demonstrate the greater effectiveness of the assessment center method as compared with traditional methods of evaluation; but as the authors point out, these results cannot be interpreted conclusively due to problems in design of the studies and failure to match adequately the assessed and non-assessed groups.

In contrast to the use of paper and pencil tests of abilities and personality traits which, when used to pick first level or even higher level managers have shown validities seldom higher than the .20s (Ghiselli, 1966), validity coefficients of assessment ratings have--with few exceptions--more often ranged from the .30s to the .60s. And several studies in AT&T and SOHIO (Bray & Grant, 1966; Bray & Campbell, 1968; Carleton, 1970; Finley, 1970) have reported that overall assessment ratings account for 10% to 20% more of the criterion variance than tests alone (Cohen, et al., 1974).

We are left, I think, with reason to believe that the assessment center method is superior to simpler and less expensive methods of discovering managerial potential, but a final judgment waits upon the completion of more definitive research.

Incidental Benefits

Aside from the question as to how effectively assessment centers meet their primary objective of effectively identifying managerial skills and potential, several surplus benefits have been thought to flow from them. One is the value of being assessed as an educational experience.

A fascinating finding (Campbell & Bray, 1967; Meyer, Note 15) is that individuals who have attended an assessment center, even those who have been evaluated as having poor potential, do better on the job than individuals who have not been assessed. Does this mean that the process of being assessed provides a participant with an opportunity of gaining a more objective picture as well as a better understanding of himself, especially in relation to others and to the requirements and demands of being a manager?

This is the testimony of most assessees (Kraut, 1972a), but self-reports of this sort do not constitute a clear demonstration of the developmental value of assessment qua assessment. Does assessment provide the candidate incidental training through participation in such exercises as the In-Basket even though there is no immediate feedback? In some centers, the In-Basket interview is used quite purposefully as a means of instructing the candidates concerning better and less good administrative practices. But even without such direction, are the assessment exercises learning experiences with beneficial consequences for later performance? Or is merely being chosen for assessment, a recognition of some promise in the candidate, motivating? Skeptics favor a quite different hypothesis; namely, that those nominated are called to the attention of important members of higher management and this exposure increases the probability of their promotion. As among the alternative hypotheses, only research can clarify the picture.

An additional benefit claimed for assessment centers is that they provide important training for line-managers who serve as assessors. In this role, the managers receive training in observing and recording behavior, and in rating relevant dimensions of managerial behavior. Observing assessees interacting in the group exercises may make the managers more aware of group dynamics and leadership styles. This experience of observing others in simulations may also give them greater appreciation of individual differences and, quite possibly, increased insights into themselves both as persons and as managers. Serving as part of an assessment team charged with formulating an estimate of the capabilities of assessees and the most efficient utilization of their

skills within the company may give the assessors an appreciation of previously unrecognized personnel problems and the manpower needs of the organization.

All of these seem likely hypotheses about the possible benefits that serving as an assessor may have upon a manager's own behavior. But until the assessors themselves are assessed, and until their subsequent records as managers are compared with that of matched samples of managers who have not served as assessors, they will remain only what they are today--namely, hypotheses.

If line-managers do become better judges of persons, more able to differentiate and to estimate the relative strengths and weaknesses of others, as a result of their serving as assessors, they may as managers be able to supply more accurate appraisals of job performance, thus contributing importantly to the improvement of performance criteria.

Possible Defects

If incidental benefits of assessment centers have been claimed without solid proof of their existence, so have critical charges been directed at them in the absence of substantiating evidence.

One of the most frequently heard charges, especially in liberal circles, is that assessment centers militate against organizational change by insuring the selection of new managers just like the old. According to this view, assessment centers reject spontaneous, imaginative, independent, non-conforming persons in favor of the "organization man."

What little evidence exists on this point suggests that this is not always the case. In one small study (N = 21) at SOHIO, the conformity score on one test showed the highest correlation of any variable with the composite criterion, namely, -.72. This offers some slim support for the investigator's conclusion that, "Apparently the 'organization man' is not greatly prized at Standard Oil Company of Ohio (Carleton, 1970, p. 566)."

Relevant to this question is the finding at IBM (Dodd, 1970) that while supervisors may nominate for assessment some candidates who are high on conformity and low on independence, the assessment center, at least in this case, did not favor this kind of person.

Put in a larger context, the question of whether the organization man is favored by the assessment center method becomes a question of whether assessment centers are being provided the opportunity to assess the right persons. The general rule is that those sent to assessment and possible promotion into management are those who are considered by their supervisors to have the greatest promise and skill. That on the average only about one-half of those sent to assessment centers are recommended for promotion suggests either that supervisors are not very good judges of managerial talent, or that they let personal biases color their judgment, or that there is little managerial potential among the rank and file of workers, or that many with promise are overlooked, possibly because they are too independent or are not sympatico with their immediate superior.

The experience of Michigan Bell is relevant. For some period of time after the initiation of self-nominations for attendance at the assessment center, self-nominees averaged higher overall ratings than did those nominated by superiors. This difference did not continue, however, suggesting that the pool of self-nominated candidates with real promise is limited but none the less worth exploring.

The pool of candidates with real potential might well be increased by casting the assessment net more widely so as to include, in addition to supervisory nominations and self-nominations, nominations by peers and from personnel records. With a wider selection of candidates to be assessed, there might well be fewer who will feel that they have been unfairly neglected. If that turned out to be the case, there might be consequent heightening of morale; but once again we lack empirical data to let us know just what the consequences are of being passed over for assessment.

Another concern is that those who receive very high ratings in assessment will be promoted, whether or not their performance on the job justifies advancement. If this were to happen, the predictions concerning the future of these candidates would become self-fulfilling prophesies. Once again, we lack firm data.

The obverse of this "crown prince or princess" effect is what has been called the "kiss of death." Here the concern is that those who do not receive high ratings in the assessment center will be more or less permanently barred from advancement, with consequent lowering of morale and motivation and possible loss to the company of these workers who may still have much needed skills. It is known that those who are promoted despite relatively poor assessment ratings tend to get the less demanding assignments; and job

opportunities and challenge have been shown to be related positively to advancement in managerial responsibility (Bray et al., 1974).

On the other hand, another study (Kraut & Scott, 1972) discovered that the percentages of low-rated and high-rated employees who left the company did not differ significantly. Perhaps we need not be too concerned, then, that assessment programs lead to a loss of well-trained, capable, and adequately performing workers, but here again we need more data.

In his later years, a major concern of H. Smith Richardson, Sr., was that his company--indeed all companies--would have too short a time perspective, recruiting and advancing those well-fitted to meet the demands of the present but ill-suited to meet the changing requirements of tomorrow. None of us has a crystal ball that will tell us with certainty what the future will be like, but there is need for all assessors to seek, from as many sources as possible, intimations of what the industrial environment is becoming, and to give serious thought concerning new assessment techniques that should be developed to meet the challenge of tapping dimensions of performance not yet dreamed of as relevant for managerial success.

With the trend toward multinational industrial organizations, perhaps more attention should be given to the candidate's sensitivity to cultural differences and his ability to live comfortably and productively in work environments that are both physically and socially quite different from anything he has known. In addition, since spouses are so important for the morale of their mates, (especially if they are destined for overseas assignment) informal assessment programs of an innovative nature might well be considered for the assessment of married couples instead of individuals.

Compositing Assessment Data

Returning again to research on assessment centers, the use of several staff members as observers in a typical assessment program, and the attempt to profit from their varying idiosyncratic observations and impressions, raises the question of how they can best be combined or reduced to a single manageable statement. Should assessors seek to reach a consensus about the overall ratings to be given candidates through the airing of all opinion, as is the custom in a final integrating staff meeting, or by a simple arithmetical averaging or compositing of independently

recorded judgments; namely, the ratings already made by the assessors in the separate assessment exercises.

In an IBM study (Wollowick & McNamara, 1969), the overall rating based on staff discussion of all the variables correlated .37 with the criterion of increase in managerial responsibility after three years. On the other hand, the multiple R based on predictors combined statistically rather than on the basis of staff discussion yielded an R of .62. This one study certainly does not answer the question of how a final overall rating is most effectively reached, since its findings have not been cross-validated, but it does point up the need for more research in this area of concern.

This same IBM study also provided some suggestive evidence that psychometric tests may be better predictors than they generally have been thought to be. The multiple R with the criterion for tests alone was .45, higher than the multiple R for either exercises alone, .39, or characteristics alone, .41. In our own study of architects (Hall & MacKinnon, 1969), we found several psychological tests to correlate significantly with the creativity of the architects as judged by architectural experts. Limiting the multiple regression equation to three scales for each test, concurrent rather than predictive correlations with the criterion, significant in both an initial sample and a cross-validating one, were for the Strong Vocational Interest Blank, .80 and .56; the California Psychological Inventory, .57 and .47; the Myers-Briggs Type Indicator, .53 and .42.

Creativity is not managerial effectiveness, nor are significant concurrent correlations as convincing as predictive validities, but still I think it is too early to give up on psychometric tests as predictors of managerial effectiveness. I would not, however, suggest that they be used alone but in conjunction with other types of assessment data, since in the IBM study as well as in other researches (Dunnette, 1971), all three types of information (tests, exercises, and characteristics) have been found to contribute importantly to the predictive success of assessment programs and none should be ignored.

Most research on assessment centers in industry has been directed toward (1) the gaining of a better understanding of the assessment process in identifying managerial skills and potential, and (2) evaluating the validity of the assessment center process. Assessment, however, is more than a method that can be applied for purposes of selection and prediction of future performance in training and subsequent careers and whose validity can be determined by appropriate research; it is also a research method for the intensive study of highly effective individuals, their development,

their structure, and their functioning. Ideally, in assessment centers there is no divorce between the use of assessment as an applied method for identifying those with promise and as a research method for the study of their lives.

Fortunately, Douglas Bray saw this in setting up the Management Progress Study, which has come to such magnificent fruition in his book Formative Years in Business, a long-term study of managerial lives. The delineation of nine life themes and changes in the relative emphasis of these themes in relation to more and to less successful careers in management, and the demonstration of their distinctive patterns of interaction in two types of managers, enlargers and enfolders, add up to a significant contribution to our understanding of the interaction of personality and career, an understanding that has many implications for the selection and placement of managers in large organizations like AT&T.

Significant research on types of managers and managerial styles need not be limited to assessment programs designed for pure research but can be carried out effectively in operational assessment programs as Jon Bentz (Note 16) has done at Sears. Not all managers will be of one type, and Bentz has impressively documented the existence of three managerial styles--those of the merchant, the operator, and the general administrator.

One can only hope that these examples of research into managerial styles and lives in progress will be followed by other assessment centers, thus realizing the full potential of assessment center methods not only to identify those with managerial potential but to add to our understanding of the factors that shape the lives of those who pursue careers in management.

Methodology

If, as the research studies seem to indicate, assessment programs are today carried out with considerable validity, how is their effectiveness to be explained? What specific factors are at work? And, if it is true that assessment centers today are, in general, more valid than the early ones, how account for that? Perhaps we can only speculate, but speculation may be of some value.

The early assessors, most often clinical psychologists and psychiatrists, had been strongly influenced by Freudian theory. As depth psychologists, they believed that behavior is strongly influenced by unconscious motives. Therefore, if one wanted to predict how a person would behave in the future, one observed his present behavior to make inferences about unconscious dynamic forces which would importantly determine what he would do in the future. The early assessors did not consider behavior in its own right but took it as a sign of something else, namely the personality structure of the individual from which inferences could be made about how he would behave in the future.

Assessors today are more inclined to view any behavior as a sample of behaviors likely to be exhibited in the future. Their more parsimonious assumption is that one will behave in the future rather much as he has behaved in the past: the best predictor of future behavior in a given situation is past behavior in a similar situation. Of course, the early assessors believed this too, otherwise they would not have developed and used simulations; but they tended, more than present day assessors, to allow their hunches as to how their assessees would behave in the future to be shaped by inferences about personality structure and its dynamics.

The use of managers as assessors has focused attention not so much on behavior in general as upon behavior in the context of management situations and demands. Managers know the situations in which managers manage and, more specifically, as they exist in their own company. Both the use of business simulations in assessment and the employment of managers as assessors reveal an implicit recognition that behavior is contingent upon the specific context and situation in which it occurs. Assessment is not, as some critics have charged, based solely upon a simplistic trait psychology but upon a broad psychology of behaviors in a social context.

Another reason why assessments may be as valid as they are is that assessors are trained to see their task as being, first and foremost, to observe precisely what assessees do in the simulation exercises and, in their report, to describe accurately what they have done; the report is a description of overt behavior. If the observer of an assessment exercise is performing as he should, he writes descriptive statements to be read to other assessors in the integration session which convey to those who have not observed the assessees a clear picture of just what the latter did. Interpretation and evaluation have little place in these reports. There is a postponing of judgment, a shunning of premature closure, a waiting to evaluate and interpret until all the data, the observations of all staff members, are in. This was not the practice, for example, in the OSS program.

Despite the widespread criticism of projective techniques as lacking demonstrated validity, they have proved more valid in assessment center practice, at least in the SOHIO FACT program (Finley, 1970), than might have been expected. This without doubt is a tribute to the interpretive skill of the projectivist but also a function, I believe, of the fact that test protocols are not used in an attempt to reveal the respondent's personality as a whole but, rather are interpreted more parsimoniously as possible indications of how the respondent behaved in the assessment exercises, and how he might be predicted to behave in a managerial role.

Obviously there are other factors contributing to the validity of assessment center ratings, but time does not permit a spelling out of all of them.

New Uses for Assessment Centers

Recent years have seen not only a wide acceptance but a broadening of the assessment center method. First used in industry to identify those with potential for moving into management, it has been extended to the assessment of those thought to possess quite different skills such as sales, technical, and entrepreneurial. Applied first at the entry level, the method is being increasingly used to assess potential for middle, higher, and even the highest levels of management.

More recently, the assessment center method has found acceptance in the field of education: for example, in graduate schools of business: at Stanford (Harrell, Note 17) to acquaint students with the techniques of assessment, and at Baylor (Williams, Note 18) to identify the particular strengths and weaknesses of students so that more effective educational programs can be tailor-made to meet their needs as future managers.

By far the most radical and innovative application of the assessment center concept in education has been made by Alverno College (Loacher, 1974; Alverno College, Note 19), a Catholic Women's College in Milwaukee, in their program variously called Liberal Learning in a Management Context, Learning in the Context of Managing One's Life, or Competence Based Learning. The program involves the student in

acquiring eight competencies or abilities:

communication skills

analytical capabilities

problem solving skills

making independent value judgments and decisions

social interactions skills

understanding the relationship between the individual and her environment

understanding of the contemporary world

knowledge and understanding of and responsiveness to the arts and humanities.

For each competence, six levels of skill are delineated and the student is supplied with detailed information as to what is expected at each level. For graduation a student must demonstrate that she has achieved the top level, 6, in at least one competence, level 5 in six others, and level 4 in the eighth competence. Competence may be evaluated by classroom performance, but more typically is assessed by a team consisting of a faculty member, a business or professional person from the Milwaukee community, an alumna, and an advanced student, who are trained for their roles as assessors. A variety of group procedures, suitable for four to six students, are employed to test the level of competence achieved by the students. Although courses are offered, credit toward graduation is not acquired by the passing of a required number of courses but by mastering the eight competencies.

Concern with the theory and practice of assessment centers is today international in scope, fostered in part by a loose organization of assessors which has already held two international congresses on the assessment center method in Williamsburg, Virginia (1973) and West Point, New York (1974). The third will be held in 1975 in Quebec City, Canada.

The congresses are only one expression of the professionalism which characterizes the field of assessment centers. It is hard to find much proprietary secrecy; assessors are open about their techniques, sharing their know-how and their knowledge with each other and with those who have a legitimate interest in their activities. To

date, research on assessment centers has come from few sources, but that appears to be changing. More research is being conducted by more centers although, since much of it is longitudinal in nature, little has so far been published.

Ethics of Assessment Center Practice

Now that procedures for conducting assessment centers and consultative services to industries are so widely available, a concern is beginning to surface among responsible assessors that unprofessional practices and unethical uses of assessment center methods which are known to have occurred in some quarters may become widespread. The last act of the 2nd International Congress on Assessment Centers was to establish a committee charged with the task of preparing a report to define what an assessment center is and to draw up guidelines for the ethical practice of assessment center methods--a professional code of ethics, if you will.

A strong trend in assessment center practice, again reflecting the professionalism of assessors, is a growing concern that all subjects be treated fairly and with respect. The increasing ethical sensitivity of assessors has doubtless been nudged by the changing climate in the United States which produced the Civil Rights Act, which in turn led to the Guidelines of the Equal Employment Opportunity Commission, but I think it is a reflection of a genuine concern by assessors that the assessment process be used in the best interests of those who are subjected to it.

There was a time when little thought was given to the role of the assessee in the assessment process; he was merely a person to be assessed by a team of assessors and then selected or rejected, promoted or not with a minimum of feedback. But today it is increasingly stressed that the assessee should be involved in his own assessment and have some input into it, if not control over it. A minimum way of involving an assessee in his own assessment and development is to provide him with a report on how he was seen by the assessment staff and by his peers.

More active and involving forms of feedback include the playing back to assessees of videotapes of their assessment exercises upon which both assessors and assessees comment. The videotape can be stopped and the assessee asked, "What could you have done differently?" Or an assessee can be shown an exceptional In-Basket with which he compares his own, focusing upon the ways in which they are similar and different.

An even more extreme form of involving an assessee in his own assessment would be to give the assessee the staff's reports on his behavior in each of the simulation exercises, and then let him, in collaboration with a member of the staff or even alone, write his own final report (Cohen, 1974b). Many would certainly question the wisdom of such practices, and the more an assessee is led to participate actively in his own assessment, the more likely it seems that the collaborating staff member should possess the special skills of a trained professional or a psychologist.

If this involvement of a participant in his own assessment were carried to the extreme, assessment centers as we have known them might no longer be felt to be necessary or perhaps even helpful. The push might then come to replace assessment centers with some such self-administered Career Development Program or Career Action Planning (Storey, 1973, 1974) which has replaced the assessment center in the General Electric Company. Valuable as such programs may be, it would be unfortunate if they were to become viewed as replacements for assessment centers. They have a place, but so, I continue to believe, do assessment centers.

Although the assessment center and the Talent Development Program of which it was a part have been discontinued in General Electric, the Career Development Program and Career Action Planning are not viewed within the Company as replacing the function of the Talent Development Program's assessment center. The tools and functions of self-directed career planning programs are different from, and not replacements for, those of assessment centers.

Further, although the Talent Development Program with its incorporated assessment center has been discontinued, other assessment centers continue to be used in General Electric and are, indeed, being increasingly employed. These are assessment centers designed to assess potential at lower levels of management, e.g., FAC, the Foreman Assessment Center for the selection of foremen, and the Managerial Assessment Center (MAC) for selecting unit level managers. While the FAC has been in use for several years, the MAC is currently in the early stages of development.

Indeed, if one can judge from present trends, my guess would be that assessment centers will continue to thrive, moving increasingly in the direction which the assessment program here in the Center for Creative Leadership has moved, as we heard yesterday, into being the first phase of an integrated program of assessment and development.

Last year as a Visiting Fellow here at CCL, it was gratifying to observe the Center in its evolving program

paralleling, if not actually in the vanguard of, the movement in assessment centers to use assessment and feedback to participants as the first step in training and development of managerial skills, as well as for the identification of strengths and weaknesses of participants.

Among some who have been most devoted to the development and improvement of assessment centers, I sense a concern about the intermeshing of assessment and development. On the one hand, fear is expressed that the integrity of the assessment center for purposes of selection will be lost if, already early in the assessment, candidates are given instruction and guidance in good management practice in, for example, the In-Basket Interview. The concern is that assessment data may be contaminated by changes already occuring during the assessment phase if developmental exercises are intermingled with assessment procedures. On the other hand, it is feared that training and development may be wide of the mark unless they are based upon a thorough assessment of each candidate's strengths and weaknesses uninfluenced by any effects of training and developmental experiences.

Certainly in any combined program there are likely to be pressures to speed up the assessment to provide more time for training. Some help in meeting this problem may lie in utilizing the best of electronic and computer devices that hold great promise for the immediate instantaneous recording of observations and rating judgments and the integration of them with other electronically scored and stored data, such as psychometric tests results.

An indication that programs combining assessment and development may well be the wave of the future in the assessment center movement is the offering of just such programs to interested firms by management companies that at first offered only assessment programs. I am thinking of the Career Development Center at Syntex and Development Dimensions, Inc. The latter offers, in conjunction with their assessment services, a full range of modules of inter-action modeling for the development of managerial skills.

And finally there is the program here at the Center for Creative Leadership, which has been a truly pioneering effort.

Donald W. MacKinnon

Biographical Note

A native of Maine, Donald W. MacKinnon received his undergraduate training at Bowdoin College and did graduate work at Harvard University, receiving the Ph.D. in psychology from Harvard in 1933. He has taught at the University of Maine, Harvard University and Radcliffe College, Bryn Mawr College, and from 1947 until his retirement in 1970 at the University of California, Berkeley.

In 1930-31 he was a Sheldon Traveling Fellow from Harvard University studying in Europe, and in the summer of 1953 he taught at the Salzburg Seminar in American Studies in Salzburg, Austria. He has also held summer session teaching appointments at the University of Minnesota (1940) and at the University of Hawaii (1955 and 1969).

During World War II he was a member of the assessment board and from June 1944 to V-J day he was director of Station S in the Office of Strategic Services.

From its founding in 1949 until his retirement in 1970, Dr. MacKinnon served as director of the Institute of Personality Assessment and Research on the Berkeley campus of the University of California. He is currently Emeritus Professor of Psychology and Research Psychologist in the Institute.

Dr. MacKinnon is the author of several scientific articles and co-author of two books: Assessment of Men, a report of the assessment program in the Office of Strategic Services during World War II, and Experimental Studies in Psychodynamics.

Dr. MacKinnon was chosen by the American Psychological Association as the Walter Van Dyke Bingham Memorial Lecturer for 1962. His lecture on "The Nature and Nurture of Creative Talent" was given in April 1962 at Yale University and has since been published in The American Psychologist. In 1967 he received, again from the American Psychological Association, the Richardson Creativity Award in recognition of his research in creativity.

During 1973-74, Dr. MacKinnon was Visiting Fellow at the Center for Creative Leadership, Greensboro, North Carolina, and Adjunct Professor of Psychology, University of North Carolina, Chapel Hill. During the year he made for the center an extensive study of the current status of assessment centers in the United States.

REFERENCES

Bender, J. M. What is "typical" of assessment centers? Personnel, 1973, 50(4), 50-57.

Bentz, V. J. The Sears experience in the investigation, description and prediction of executive behavior. In F. R. Wickert & D. E. McFarland (Eds.), Measuring executive effectiveness. New York: Appleton-Century-Crofts, 1967.

Bray, D. W. The assessment center method of appraising management potential. In J. W. Blood (Ed.), The personnel job in a changing world. New York: American Management Association, 1964. (a)

Bray, D. W. The management progress study. American Psychologist, 1964, 19, 419-420. (b)

Bray, D. W., & Campbell, R. J. Selection of salesmen by means of an assessment center. Journal of Applied Psychology, 1968, 52, 36-41.

Bray, D. W., Campbell, R. J., & Grant, D. L. Formative years in business: A long-term AT&T study of managerial lives. New York: Wiley-Interscience, 1974.

Bray, D. W., & Grant, D. L. The assessment center in the measurement of potential for business management. Psychological Monographs, 1966, 80(17, Whole No. 625).

Bray, D. W., & Moses, J. L. Personnel selection. In P. H. Mussen & M. R. Rosenzweig (Eds.), Annual review of psychology (Vol. 23). Palo Alto, Calif.: Annual Reviews, 1972.

Byham, W. C. Assessment centers for spotting future managers. Harvard Business Review, 1970, 48(4), 150-160, plus appendix.

Campbell, R. J., & Bray, D. W. Assessment centers: An aid in management selection. Personnel Administration, 1967, 30(2), 6-13.

Carleton, F. O. Relationships between follow-up evaluations and information developed in a management assessment center. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 565-566. (Summary)

Cohen, B. M. The assessment center: Whom to develop. Training in Business and Industry, February 1974, pp. 19-21. (a)

Cohen, B. M. Increased learning opportunities from the assessment center technique. Assessment & Development, 1974, 2(1), 4. (b)

Cohen, B. M., Moses, J. L., & Byham, W. C. The validity of assessment centers: A literature review. Monograph II. Pittsburg, Pa.: Development Dimensions, 1974.

Cronbach, L. J. Assessment, 1955. Presidential Address, Division 5, Annual Meeting, American Psychological Association, San Francisco, September 1955. American Psychologist, 1955, 10, 419. (Reference)

Cronbach, L. J. Assessment of individual differences. In P. R. Farnsworth & Q. McNemar (Eds.), Annual review of psychology (Vol. 7). Palo Alto, Calif.: Annual Reviews, 1956.

Cutterbuck, D. Acid test for management potential. International Management, 1974, 29, 54-57.

Development Dimensions. Catalog of assessment and development exercises. Pittsburgh, Pa.: Development Dimensions, 1975.

Dodd, W. E. Will management assessment centers insure selection of the same old types? Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 569-570. (Summary)

Donaldson, R. J. Validation of the internal characteristics of an industrial assessment center program using the multitrait-multimethod matrix approach. Unpublished doctoral dissertation, Case Western Reserve University, 1969.

Dunnette, M. D. Multiple assessment procedures in identifying and developing managerial talent. In P. McReynolds (Ed.), Advances in psychological assessment (Vol. 2). Palo Alto, Calif.: Science and Behavior Books, 1971.

Finkle, R. B., & Jones, W. S. Assessing corporate talent: A key to managerial manpower planning. New York: Wiley-Interscience, 1970.

Finley, R. M., Jr. Evaluation of behavior predictions from projective tests given in a management assessment center. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 567-568. (Summary)

Garforth, F. I. de la P. War Office Selections Boards. Occupational Psychology, 1945, 19, 97-108.

Ghiselli, E. E. The validity of occupational aptitude tests. New York: Wiley-Interscience, 1966.

Greenwood, J. M., & McNamara, W. J. Interrater reliability in situational tests. Journal of Applied Psychology, 1967, 31, 101-106.

Hall, W. B., & MacKinnon, D. W. Personality inventory correlates of creativity among architects. Journal of Applied Psychology, 1969, 53, 322-326.

Hinrichs, J. R. Comparison of "real life" assessments of management potential with situational exercises, paper-and-pencil ability tests, and personality inventories. Journal of Applied Psychology, 1969, 53, 425-433.

Holt, R. R., & Luborsky, L. Personality patterns of psychiatrists. New York: Basic Books, 1958.

Howard, A. An assessment of assessment centers. Academy of Management Journal, 1974, 17, 115-134.

Huck, J. R. Assessment centers: A review of external and internal validities. Personnel Psychology, 1973, 26, 191-212.

Jaffee, C., Bender, J., & Calvert, D. The assessment center technique: A validation study. Management of Personnel Quarterly, 1970, 9, 9-14.

Kelly, E. L., & Fiske, D. W. Prediction of success in the V. A. training program in clinical psychology. American Psychologist, 1950, 5, 395-406.

Kelly, E. L., & Fiske, D. W. The prediction of performance in clinical psychology. Ann Arbor: University of Michigan Press, 1951.

Kelly, E. L., & Goldberg, L. R. Correlates of later performances and specialization in psychology. Psychological Monographs, 1959, 73(12, Whole No. 482).

Kraut, A. I. Intellectual ability and promotional success among high level managers. Personnel Psychology, 1969, 22, 281-290.

Kraut, A. I. A hard look at management assessment centers and their future. Personnel Journal, 1972, 51, 317-326. (a)

Kraut, A. I. The validity of an operational management program. Journal of Applied Psychology, 1972, 56, 124-129. (b)

Kraut, A. I., & Scott, G. J. Validity of an operational management assessment program. Journal of Applied Psychology, 1972, 56, 124-129.

Loacher, G. Assessment procedures used as educational criterion. Assessment & Development, 1974, 2(1), 5.

MacKinnon, D. W. The use of clinical methods in social psychology. Journal of Social Issues, 1946, 2, 1-8.

MacKinnon, D. W. Some problems of assessment. Transactions of the New York Academy of Science, 1947, Ser. II, 9, 171-185.

MacKinnon, D. W. The highly effective individual. Teachers College Record, 1960, 61, 367-378.

McConnell, J. J., & Parker, T. C. An assessment center program for multi-organizational use. Training and Development Journal, 1972, 26(3) 6-14.

Miller, D. Career development: It's your job but you need help. Chemical Engineering, 1974, 81(12), 94-98.

Murray, H. A., et al. Explorations in personality. New York: Oxford University Press, 1938.

Murray, H. A., & MacKinnon, D. W. Assessment of OSS personnel. Journal of Consulting Psychology, 1946, 10, 76-80.

OSS Assessment Staff. The assessment of men. New York: Rinehart, 1948.

Simoneit, M. Deutsches Soldatentum, 1914 und 1939. Berlin: Junker und Buennhaupt, 1940.

Simoneit, M. Grundriss der Charakterologischen Diagnostik. Leipzig: Teubner, 1944.

Storey, W. D. How to develop a career plan: Basic theory and methodology for self-directed career planning. Proceedings of the Third Annual Frontiers for Education Conference. Layfayette, Ind.: Purdue University, 1973.

Storey, W. D. An introduction to self-directed career planning. In S. S. Dubin, H. Shelton, & J. McConnell (Eds.), Maintaining professional and technical competence of the older engineer--engineering and psychological aspects. Washington, D. C.: American Society for Engineering Education, 1974.

Thomson, H. A. Internal and external validation of an industrial assessment center program. Unpublished doctoral dissertation, Case Western Reserve University, 1969.

Thomson, H. A. A comparison of predictor and criterion judgments of managerial performance using the multitrait-multimethod approach. Journal of Applied Psychology, 1970, 54, 496-502.

Wollowick, H. B., & McNamara, W. J. Relationships of the components of an assessment center to management success. Journal of Applied Psychology, 1969, 53, 348-352.

REFERENCE NOTES

1. Finkle, R. B. Managerial assessment centers. Book in preparation, 1974.

2. Hinrichs, J. R., & Haanperä, S. A technical research on management assessment in the IBM World Trade Corporation (Personnel Research Study No. 18). IBM Europe, March 1974. (Abstract)

3. Bentz, V. J. In-basket factors, combined years: 1968 and 1969. Chicago: Sears Roebuck & Co., August 1969.

4. Moses, J. L. Assessment center performance and management progress. In W. C. Byham (Chair), Validity of assessment centers. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

5. Dodd, W. E. Validity studies at IBM. In W. C. Byham (Chair) Validity of assessment centers. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

6. Bentz, V. J. Sears longitudinal study of management behavior. In A. C. MacKinney (Chair), Longitudinal approaches to management research. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

7. Bentz, V. J. Supporting data for the Sears longitudinal study of management behavior. In A. C. MacKinney (Chair), Longitudinal approaches to management research. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

8. Bentz, V. J. Tables supporting presentation of validity of Sears assessment center procedures. In W. C. Byham (Chair), Validity of assessment centers. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

9. Bentz, V. J. Validity of Sears assessment center procedures. In W. C. Byham (Chair), Validity of assessment centers. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

10. Bentz, V. J. Tables supporting presentation of validity of Sears assessment center procedures. In W. C. Byham (Chair), Validity of assessment centers. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

11. Bullard, J. F. An evaluation of the assessment center approach to selecting supervisors. Peoria, Ill.: Caterpillar Tractor Company, Corporate Personnel Department, May 1969.

12. Michigan Bell Telephone. Personnel assessment program: Evaluation study. Detroit: Michigan Bell Telephone, 1962.

13. Meyer, H. H. Assessment centers at General Electric. Paper presented at Development Dimensions Orientation Conference, San Francisco, March 1972.

14. Bullard, J. F. An evaluation of the assessment center approach to selecting supervisors. Peoria, Ill.: Caterpillar Tractor Company, Corporate Personnel Department, May 1969.

15. Meyer, H. H. Assessment centers at General Electric. Paper presented at Development Dimensions Orientation Conference, San Francisco, March 1972.

16. Bentz, V. J. A basic investigation into the nature of managerial styles. Chicago: Sears Roebuck & Co., 1974.

17. Harrell, T. B. 371--Career planning, Spring 1974. Graduate School of Business, Stanford University, 1974.

18. Williams, J. C. Manager education--a personnel approach. Waco, Tex.: Baylor University, 1974. (Abstract)

19. Alverno College. Competence based learning at Alverno College. Milwaukee: Alverno College, 1974.